

## СТРАТЕГИЯ КОДИРОВАНИЯ АЛКОГОЛЬДЕГИДРОГЕНАЗ ЧЕЛОВЕКА

БУТВИЛОВСКИЙ А.В.\* , БАРКОВСКИЙ Е.В.\*\* ,  
БУТВИЛОВСКИЙ В.Э.\*\*\*

*УО «Белорусский государственный медицинский университет»;  
центральная научно-исследовательская лаборатория\* ,  
кафедра общей химии БГМУ\*\* ,  
кафедра биологии БГМУ\*\*\**

**Резюме.** Статья посвящена определению стратегии кодирования алкогольдегидрогеназ человека. Установлено наличие АТ-давления в изучаемых последовательностях мРНК, кодирующих АДГ классов 1А, 1С, 2, 3, 4, 5 и слабого ГЦ-давления в последовательности мРНК, кодирующей АДГ 1В класса. Стратегия кодирования алкогольдегидрогеназ человека заключается в том, что при слабом ГЦ-давлении по сравнению с АТ-давлением преимущественно используются синонимичные ГЦЗ-кодоны, кодирующие изолейцин, серин, треонин, гистидин, глутамин, аспарагин, аргинин и глицин, увеличивается использование ГЦЗ-кодонов, кодирующих лейцин, валин, пролин, аланин тирозин, лизин и глутаминовую кислоту, а также уменьшается использования ГЦЗ-кодонов, соответствующих фенилаланину, цистеину и аспарагиновой кислоте.

**Ключевые слова:** стратегия кодирования, предпочтение кодонов, алкогольдегидрогеназа.

**Abstract.** This article is devoted to definition of the coding strategy for human alcohol dehydrogenases. We established presence of AT-pressure in mRNA sequences, coding ADH 1A, 1C, 2, 3, 4, 5 classes and weak GC-pressure in mRNA sequence, coding ADH 1B class. The coding strategy for human alcohol dehydrogenases consists that at a weak GC-pressure in comparison with AT-pressure synonymous GC3-codons of ile, ser, thr, his, gln, asn, arg, gly will mainly be utilized, usage of GC3-codons coding leu, val, pro, ala, tyr, lys and glu is enlarged and usage of GC3-codons appropriate to phe, cys and asp is decreased.

**Адрес для корреспонденции:** Республика Беларусь, 220116, г. Минск, пр Дзержинского, 83, т. (017) 272-65-98, (0295) 65-50-00,  
Бутвиловский А.В.

Из-за вырожденности генетического кода большинство аминокислот кодируются более чем одним кодоном. Так как синонимичные замены не вызывают изменения кодируемой аминокислоты, а естественный отбор действует главным образом на уровне белков, то синонимичные замены следует считать нейтральными. Если это так, то частоты использования кодонов в пределах одной серии должны быть одинаковыми. Однако зачастую это не так, то есть существует предпочтение кодонов – явление, при котором частоты использования кодонов в серии различаются.

Для объяснения этого явления предложено ряд гипотез. Так в 1979 году Л. Пост связал предпочтение кодонов с количеством соответствующих тРНК и со степенью экспрессии данного гена. В последующем установлено наличие очень сильной корреляции между содержанием различных тРНК и частотой использования соответствующих кодонов в генах дрожжей и *E.coli*. Однако может иметь место и обратная ситуация, когда соотношение между различными тРНК, определяемое числом их генов, подгоняется к относительным частотам синонимичных кодонов. Вполне вероятно, что количество изоакцепторных тРНК и значения RSCU взаимосвязаны, то есть изменение одного из этих показателей вызывает изменение другого и наоборот. Также были высказаны гипотезы о влиянии на неслучайное использование синонимичных кодонов содержания гуанина и цитозина, а также о возможной роли wobbling (способности тРНК узнавать не только свой кодон) особенно в случае модификации оснований в первом положении антикодона. Следует сказать, что в начале 80-х годов прошлого столетия Р. Грэнтсем предложил геномную гипотезу, согласно которой большинство генов данного генома имеют одинаковую кодирующую стратегию [3].

В предварительно проведенных исследованиях нами обнаружено наличие влияния ГЦ-давления на использование претерминальных кодонов в последовательностях мРНК, кодирующих алкогольдегидрогеназы (АДГ) класса 3 [4]. По этой причине, с нашей точки зрения, содержание гуанина и цитозина является одним из наиболее вероятных факторов, определяющих неслучайное использование синонимичных кодонов.

Цель исследования: установить стратегию кодирования алкогольдегидрогеназ человека.

#### Методы

Проанализированы аминокислотные последовательности алкогольдегидрогеназ человека классов 1A, 1B, 1C [8], 2 [7], 3 [6], 4 [5], 5 [11], а также кодирующие их нуклеотидные последовательности мРНК. Для определения аминокислотного и нуклеотидного состава, показателя относительного использования синонимичных кодонов (RSCU, relative synonymous codon usage) использовался пакет программ MEGA 3 [9]. Полученные результаты обработаны методами описательной статистики, достоверность различий определена по критерию t.

#### Результаты

Как известно, молекула ДНК состоит из двух антипараллельных полинуклеотидных цепей, соединенных друг с другом при помощи водородных связей, возникающих между азотистыми основаниями разных цепочек. Причем

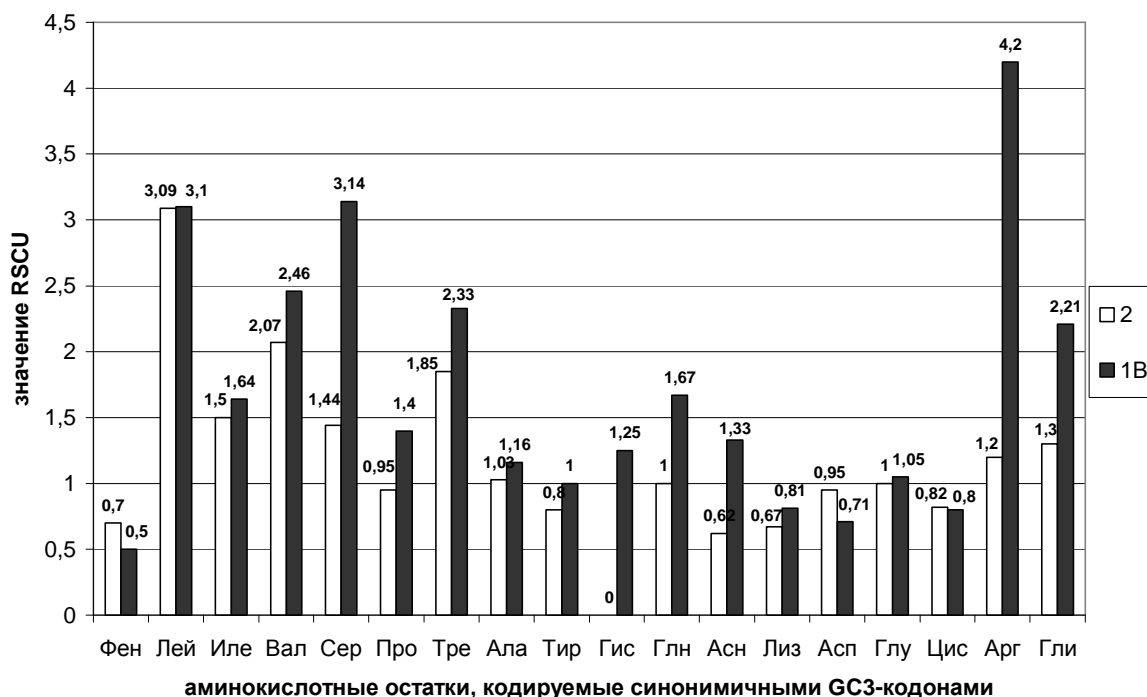
между аденином и тиминном устанавливаются две связи, а между гуанином и цитозином – три. Насыщенность молекулы ДНК гуанином и цитозином является одним из факторов, обеспечивающих ее термодинамическую стабильность. Об этом можно судить исходя из насыщенности гуанином и цитозином мРНК, образовавшейся в процессе транскрипции. Для определения стратегии кодирования алкогольдегидрогеназ человека установлено процентное содержание гуанина и цитозина в соответствующих последовательностях мРНК (табл. 1).

Таблица 1

**Процентное содержание гуанина и цитозина в последовательностях мРНК, кодирующих алкогольдегидрогеназы человека**

Класс АДГ	G+C, %	G1+C1, %	G2+C2, %	G3+C3, %
1A	47,7	53,2	42,6	47,0
1B	49,4	54,8	43,1	50,5
1C	48,2	54,0	43,0	47,6
2	44,8	50,9	44,3	39,2
3	45,9	53,6	43,2	40,8
4	46,8	51,4	43,5	45,3
5	45,6	51,0	42,1	43,9

Для установления стратегии кодирования определен показатель RSCU для кодонов, содержащих в третьем положении гуанин или цитозин, в последовательностях АДГ 1В и 2 классов человека, как наиболее отличающихся по содержанию гуанина и цитозина (диаграмма 1).



**Диаграмма 1. Показатель относительного использования синонимичных ГЦ3-кодонов для алкогольдегидрогеназ 1В и 2 классов человека.**

Для статистического подтверждения полученной зависимости между ГЦ-содержанием и значениями RSCU для синонимичных ГЦЗ-кодонов нами вычислены коэффициенты корреляции, их ошибка и достоверность корреляционных связей (табл. 2).

Таблица 2

**Показатели корреляции между содержанием гуанина и цитозина и значениями RSCU для ГЦЗ-кодонов в последовательностях мРНК, кодирующих алкогольдегидрогеназы человека**

Аминокислота	Фен	Вал	Арг	Лей	Сер	Про	Тре	Ала	Гли
GC3-кодоны	УУЦ	ГУЦ ГУГ	ЦГЦ ЦГГ АГГ	УУГ ЦУЦ ЦУГ	УЦЦ УЦГ АГЦ	ЦЦЦ ЦЦГ	АЦЦ АЦГ	ГЦЦ ГЦГ	ГГЦ ГГГ
$r \pm m$	$-0,26 \pm 0,431$	$0,76 \pm 0,291^*$	$0,58 \pm 0,363$	$-0,40 \pm 0,409$	$0,92 \pm 0,170^*$	$0,73 \pm 0,307^*$	$0,65 \pm 0,341$	$-0,38 \pm 0,414$	$0,82 \pm 0,259^*$
аминокислота	Лиз	Асн	Глн	Асп	Гис	Глу	Иле	Цис	Тир
GC3-кодоны	ААГ	ААЦ	ЦАГ	ГАЦ	ЦАЦ	ГАГ	АУЦ	УГЦ	УАЦ
$r \pm m$	$-0,16 \pm 0,441$	$0,67 \pm 0,334^*$	$0,59 \pm 0,362$	$-0,17 \pm 0,441$	$0,54 \pm 0,376$	$0,36 \pm 0,417$	$0,25 \pm 0,433$	$-0,37 \pm 0,416$	$0,62 \pm 0,350$

Примечание. Достоверные корреляционные связи ( $p < 0,05$ ) обозначены знаком \*.

В табл. 3 представлена зависимость между ГЦ-содержанием и содержанием аминокислот в изучаемых последовательностях.

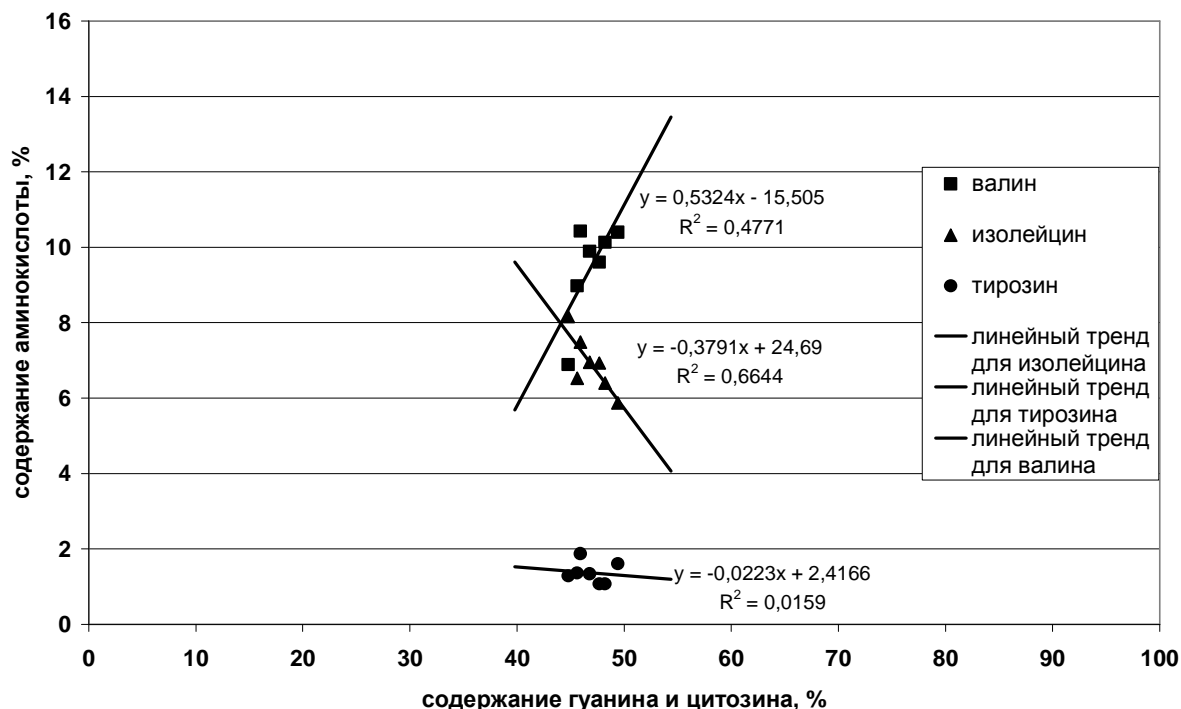
Таблица 3

**Коэффициенты корреляции между содержанием гуанина и цитозина и содержанием аминокислот в последовательностях алкогольдегидрогеназ человека**

Аминокислота	Фен	Вал	Арг	Лей	Сер	Про	Тре	Ала	Гли	Мет
$r \pm m$	$-0,26 \pm 0,431$	$0,69 \pm 0,323^*$	$0,46 \pm 0,397$	$-0,03 \pm 0,447$	$-0,44 \pm 0,402$	$0,85 \pm 0,238^*$	$-0,15 \pm 0,442$	$-0,21 \pm 0,437$	$-0,50 \pm 0,386$	$0,51 \pm 0,383$
аминокислота	Лиз	Асн	Глн	Асп	Гис	Глу	Иле	Цис	Тир	Три
$r \pm m$	$0,27 \pm 0,431$	$0,19 \pm 0,439$	$-0,48 \pm 0,393$	$0,00 \pm 0,447$	$0,65 \pm 0,338$	$-0,17 \pm 0,440$	$-0,82 \pm 0,259^*$	$-0,43 \pm 0,403$	$-0,13 \pm 0,444$	$-0,70 \pm 0,320^*$

Примечание. Достоверные корреляционные связи ( $p < 0,05$ ) обозначены знаком \*.

На диаграмме 2 показана зависимость содержания тирозина, валина и изолейцина в последовательностях алкогольдегидрогеназ человека от общего ГЦ-содержания.



**Диаграмма 2. Зависимость содержания изолейцина, валина и тирозина в последовательностях алкогольдегидрогеназ человека от общего содержания гуанина и цитозина в кодирующих их последовательностях мРНК.**

### Обсуждение

Из приведенных в табл. 1 данных видно, что общее содержание гуанина и цитозина в изучаемых последовательностях мРНК колеблется в пределах 44,8 – 49,4%, при чем наибольшее содержание вышеназванных азотистых оснований характерно для АДГ класса 1В (49,4%), а наименьшее – для АДГ класса 2 (44,8%). Большая насыщенность гуанином и цитозином последовательности мРНК, кодирующей алкогольдегидрогеназы класса 1В человека, может свидетельствовать о большей термодинамической стабильности соответствующего участка молекулы ДНК.

Кроме этого, проанализировано содержание гуанина и цитозина в каждом положении нуклеотида в кодоне в отдельности. Так, для АДГ 1А класса наибольшее содержание Г+Ц отмечено в первом положении нуклеотида в кодоне (53,2%), меньшее в третьем (47,0%), а наименьшее – во втором (42,6%). Для последовательностей мРНК, соответствующих АДГ 1А, 1С, 2, 3, 4, 5 классов, характерно преобладание общего ГЦ-содержания над содержанием гуанина и цитозина в третьем положении кодона, что позволяет говорить о наличии в изучаемых последовательностях АТ-давления. Для последовательности мРНК, кодирующей АДГ 1В класса, характерно незначительное преобладание содержания гуанина и цитозина в третьем положении кодона над общим ГЦ-содержанием, а, следовательно, наличие слабого ГЦ-давления.

На диаграмме 1 видно, что для 15 (83,3%) аминокислотных остатков, кодируемых ГЦЗ-кодонами, значение RSCU увеличивается с ростом общего со-

держания гуанина и цитозина и лишь для трех аминокислот (фенилаланина, цистеина и аспарагиновой кислоты) – уменьшается. Таким образом, при уменьшении АТ-содержания наблюдается преимущественное использование синонимичных ГЦЗ-кодонов, соответствующих 8 аминокислотным остаткам (изолейцину, серину, треонину, гистидину, глутамину, аспарагину, аргинину и глицину). При росте ГЦ-содержания происходит увеличение использования ГЦЗ-кодонов, кодирующих 7 аминокислот (лейцин, валин, пролин, аланин, тирозин, лизин и глутаминовую кислоту), а также уменьшение использования ГЦЗ-кодонов, соответствующих 3 аминокислотам (фенилаланину, цистеину и аспарагиновой кислоте). Поскольку АДГ классов 1В и 2 в основном синтезируются в одних и тех же клетках (гепатоцитах), обладающих одним пулом тРНК, трудно согласиться с мнением о том, что содержание изоакцепторных тРНК является главным фактором, определяющим предпочтительное использование синонимичных кодонов [3].

Только для пяти (валин, серин, пролин, глицин и аспарагин) из восемнадцати рассмотренных аминокислотных остатков, кодируемых GC3-кодонами, характерно наличие достоверных прямых корреляционных связей между значением RSCU и общим содержанием гуанина и цитозина. При этом для аспарагина связь средней силы, а для валина, серина, пролина и глицина – связь сильная. Между сравниваемыми показателями для фенилаланина, лейцина, аланина, лизина, аспарагиновой кислоты и цистеина характерны отрицательные корреляционные связи, не являющиеся, однако достоверными.

Изменение значения RSCU можно объяснить исходя из двух механизмов. Во-первых, изменение RSCU может быть связано с заменами в третьем положении кодона, то есть с взаимными переходами АТЗ- и ГЦЗ-кодонов. Во-вторых, изменение RSCU может быть обусловлено заменами в первом и втором положениях кодона, что приводит к изменению кодируемой аминокислоты. При наличии прямой корреляционной связи между RSCU и ГЦ-содержанием наиболее вероятен первый механизм, а в обратном случае – второй механизм. Замена кодируемой аминокислоты в этом случае позволяет предположить ее большее структурно-функциональное значение.

Таким образом, используя этот подход, можно предположить существенное значение фенилаланина, лейцина, аланина, лизина, аспарагиновой кислоты и цистеина для активности алкогольдегидрогеназ человека. Известно, что все вышеназванные аминокислотные остатки действительно необходимы для структуры и функции АДГ. Так, фенилаланин участвует в формировании внутренней части, а лейцин – наружной части субстрат-связывающего пакета, что во многом определяет субстратную специфичность алкогольдегидрогеназ. Лизин вместе с аспарагиновой кислотой играет важную роль в образовании фермент-коферментного комплекса. Остатки цистеина и аспарагиновой кислоты обеспечивают образование холофермент-субстратного комплекса, удерживаемого ионами цинка. Остатки аланина, по-видимому, принимают участие в поддержании структуры фермента.

Между ГЦ-содержанием и содержанием валина, пролина, изолейцина и триптофана существуют достоверные корреляционные связи. При чем для ва-

лина характерна прямая связь средней силы, для пролина – прямая сильная связь, а для изолейцина и триптофана – обратные сильные связи.

По показателям корреляции, полученным для алкогольдегидрогеназ человека, все входящие в их последовательности аминокислоты можно разделить на три группы. К первой группе отнесены аминокислоты, содержание которых достоверно положительно коррелирует с содержанием гуанина и цитозина (валин и пролин). Ко второй группе – аминокислоты, содержание которых достоверно отрицательно коррелирует с содержанием гуанина и цитозина (изолейцин и триптофан). А к третьей группе – аминокислоты, содержание которых не коррелирует достоверно с содержанием гуанина и цитозина (фенилаланин, аргинин, лейцин, серин, треонин, аланин, глицин, метионин, лизин, аспарагиновая и глутаминовая кислоты, глутамин, аспарагин, гистидин, цистеин и тирозин).

Наблюдаемый на диаграмме 2 перекрест трендов для валина и изолейцина можно объяснить транзициями А→Г в первом положении кодона под влиянием ГЦ-давления. В пользу этого говорит и тот факт, что транзиции более вероятны, чем трансверсии, а замены А→Г происходят чаще, чем любые другие замены [1, 2, 3].

Следует отметить, что сходная классификация аминокислот, входящих в состав бактериальных белков, в зависимости от насыщенности гуанином и цитозином кодирующих их была ранее произведена Н. Суюекой [10]. При этом 11 из 18 изученных им аминокислот (61,1%) аналогично классифицированы нами. Отсутствие полной идентичности, вероятно, связано с различиями стратегий кодирования белков у человека и бактерий, а также малым диапазоном колебаний значений ГЦ-содержания АДГ человека (44,8–49,4%).

#### Заключение

Наибольшее содержание гуанина и цитозина характерно для последовательностей мРНК, кодирующих АДГ класса 1В (49,4%), а наименьшее – АДГ 2 класса (44,8%).

Для пяти (валина, серина, пролина, глицина и аспарагина) из 18 рассмотренных аминокислотных остатков, кодируемых ГЦ3-кодонами, характерно наличие достоверных прямых корреляционных связей между значением RSCU и общим содержанием гуанина и цитозина.

По показателям корреляции между содержанием аминокислот в последовательностях алкогольдегидрогеназ человека и ГЦ-насыщенностью кодирующих их последовательностей мРНК, все аминокислоты можно разделить на 3 группы:

- 1) Аминокислоты, содержание которых достоверно положительно коррелирует с содержанием гуанина и цитозина (валин и пролин);
- 2) Аминокислоты, содержание которых достоверно отрицательно коррелирует с содержанием гуанина и цитозина (изолейцин и триптофан);
- 3) Аминокислоты, содержание которых не коррелирует достоверно с содержанием гуанина и цитозина (фенилаланин, аргинин, лейцин, серин, треонин, аланин, глицин, метионин, лизин, аспарагиновая и глутаминовая кислоты, глутамин, аспарагин, гистидин, цистеин и тирозин).

Предложенная классификация аминокислот на 61,1% аналогична классификации аминокислот входящих в состав бактериальных белков, предложенной Н. Суеоккой. Из этого следует, что для бактериальных белков и белков человека (алкогольдегидрогеназ) характерны сходные стратегии кодирования в зависимости от содержания гуанина и цитозина, несмотря на малый диапазон колебаний значений ГЦ-содержания в мРНК, кодирующих АДГ человека.

Таким образом, стратегия кодирования алкогольдегидрогеназ человека заключается в том, что при слабом ГЦ-давлении по сравнению с АТ-давлением преимущественно используются синонимичные ГЦЗ-кодона, кодирующие изолейцин, серин, треонин, гистидин, глутамин, аспарагин, аргинин и глицин, увеличивается использование ГЦЗ-кодонов, кодирующих лейцин, валин, пролин, аланин тирозин, лизин и глутаминовую кислоту, а также уменьшается использования ГЦЗ-кодонов, соответствующих фенилаланину, цистеину и аспарагиновой кислоте.

### Литература

1. Барковский Е.В., Ачинович О.В., Барабаш С.Г. [и др.] // Белорусский медицинский журнал. – 2002. – № 2. – С. 49-52.
2. Бутвиловский, А. В. // Белорусский медицинский журнал. – 2005. – №2. – С.30-32.
3. Кимура, М. Молекулярная эволюция: теория нейтральности / М. Кимура – М., 1985. – 398 с.
4. Хрусталеv, В.В. Молодежная наука и современность // Юбилейная межвузовская научная конференция студентов и молодых ученых, посвященная 70-летию КГМУ: в 2-х ч./ В.В. Хрусталеv, А.В. Бутвиловский. – Курск: КГМУ, 2005. – Ч. I. – С. - 81-82.
5. Farres J., Moreno A., Crosas B. [et al.] // Eur. J. Biochem. – 1994. – Vol. 224(2). – P. 549-557.
6. Galter D., Carmine A., Buervenich S. [et al.] // Eur. J. Biochem. – 2003. – Vol. 270 (6). – P. 1316-1326.
7. Iida A, Saito S., Sekine A. [et al.] // J. Hum. Genet. – 2002. – Vol. 47 (2). – P. 74-76.
8. Ikuta T., Szeto S., Yoshida A. // Biochemistry. – 1986. – Vol. 25 (9). – P. 2465-2470.
9. Kumar S., Tamura K., Nei M. // Brief. Bioinformatics. - 2004. - Vol. 5. – P. 150-160.
10. Sueoka N. // PNAS – 1961. – Vol. 47. – P. 1141-1149.
11. Zhi X., Chan E.M. and Edenberg H.J. // J. DNA Cell Biol. – 2000. – Vol. 19(8). – P. 487-497.